Domanda **1**

Risposta non ancora data

Punteggio max.: 1,00

Which of the following is not a property of a *metric* distance function

Scegli un'alternativa:

- a. Symmetry
- b. Triangle inequality
- c. Boundedness
- d. Positive definiteness

Annulla la scelta

Domanda **2**

Risposta non ancora data

Punteggio max.: 1,00

Which of the following *is not* an objective of feature selection

Scegli un'alternativa:

- a. Avoid the *curse of dimensionality*
- b. Reduce the effect of noise
- c. Select the features with higher range, which have more influence on the computations
- d. Reduce time and memory complexity of the mining algorithms

Annulla la scelta

Domanda **3**

Risposta non ancora data

Punteggio max.: 1,00

What is the *single linkage?*

Scegli un'alternativa:

- a. A method to compute the separation of the objects inside a cluster
- b. A method to compute the distance between two objects, it can be used in hierarchical clustering
- c. A method to compute the distance between two sets of items, it can be used in hierarchical clustering
- d. A method to compute the distance between two classes, it can be used in decision trees

Annulla la scelta

Domanda **4**

Risposta non ancora data

Punteggio max.: 1,00

Which of the following measure can be used as an alternative to the *Information Gain?*

Scegli un'alternativa:

- a. Rand Index
- b. Silhouette Index
- c. Jaccard Index
- d. Gini Index

Annulla la scelta

Domanda **5**

Risposta non ancora data

Punteggio max.: 1,00

What is the *Gini Index?*

Scegli un'alternativa:

- a. A measure of the *entropy* of a dataset
- b. An impurity measure of a dataset alternative to the *Information Gain* and to the *Misclassification Index*
- c. An impurity measure of a dataset alternative to *overfitting* and *underfitting*
- d. An accuracy measure of a dataset alternative to the *Information Gain* and to the *Misclassification Index*

Annulla la scelta

Domanda **6**

Risposta non ancora data

Punteggio max.: 1,00

A Decision Tree is...

Scegli un'alternativa:

- a. A tree-structured plan of tests on multiple attributes to forecast the target
- b. A tree-structured plan of tests on single attributes to obtain the maximum purity of a node
- c. A tree-structured plan of tests on single attributes to forecast the cluster
- d. A tree-structured plan of tests on single attributes to forecast the target

Annulla la scelta

Domanda **7**

Risposta non ancora data

Punteggio max.: 1,00

When training a neural network, what is the *learning rate*?

- a. The ratio between the size of the hidden layer and the input layer of the network
- b. A multiplying factor of the correction to be applied to the connection weights
- c. The slope of the activation function in a specific node
- d. The speed of convergence to a stable solution during the learning process

Annulla la scelta

Domanda **8**

Risposta non ancora data

Punteggio max.: 1,00

Which of the statements below about *Hierarchical Agglomerative Clustering* is true?

- a. Is very efficient, also with large datasets
- b. Is based on a well founded statistical model
- c. Requires the definition of *distance between sets of objects*
- d. Requires the definition of *Inertia* of clusters

Annulla la scelta

Domanda **9**

Risposta non ancora data

Punteggio max.: 1,00

What does K-means try to minimise?

Scegli un'alternativa:

- a. The *separation*, that is the sum of the squared distances of each cluster centroid with respect tho the global centroid of the dataset
- b. The *distortion*, that is the sum of the squared distances of each point with respect to the points of the other clusters
- c. The *separation*, that is the sum of the squared distances of each point with respect to its centroid
- d. The *distortion*, that is the sum of the squared distances of each point with respect to its centroid

Annulla la scelta

Domanda **10**

Risposta non ancora data

Punteggio max.: 1,00

Which of the following characteristic of data can reduce the effectiveness of DBSCAN?

Scegli un'alternativa:

- a. All the variables are the same range of values
- b. Presence of clusters with different densities
- c. Presence of outliers
- d. Clusters have concavities

Annulla la scelta

Domanda **11**

Risposta non ancora data

Punteggio max.: 1,00

Which of the following statements regarding the discovery of association rules is true? (One or more)

Scegli una o più alternative:

- [x] a. The confidence of a rule can be computed starting from the supports of itemsets
- [ ] b. The support of a rule can be computed given the confidence of the rule
- [x] c. The support of an itemset is anti-monotonic with respect to the composition of the itemset
- [ ] d. The confidence of an itemset is anti-monotonic with respect to the composition of the itemset

Domanda **12**

Risposta non ancora data

Punteggio max.: 1,00

How does *pruning* work when generating frequent itemsets?

Scegli un'alternativa:

- a. If an itemset is not frequent, then none of its supersets can be frequent, therefore the frequencies of the supersets are not evaluated
- b. If an itemset is not frequent, then none of its subsets can be frequent, therefore the frequencies of the subsets are not evaluated
- c. If an itemset is frequent, then none of its subsets can be frequent, therefore the frequencies of the subsets are not evaluated
- d. If an itemset is frequent, then none of its supersets can be frequent, therefore the frequencies of the supersets are not evaluated

Annulla la scelta

Domanda **13**

Risposta non ancora data

Punteggio max.: 1,00

In *feature selection*, what is the Principal Component Analysis?

Vai a...

Machine Learning - Python Lab ►

- a. A mathematical technique used to transform a set of numeric attributes into a smaller set of numeric attributes which capture most of the variability in data
- b. A mathematical technique used to transform non numeric attributes into numeric attributes
- c. A heuristic technique used to find a subset of the attributes which produces the same classifier
- d. A mathematical technique used to find the principal attributes which determine the classification process

Annulla la scelta

Domanda **14**

Risposta non ancora data

Punteggio max.: 1,00

What is the difference between classification and regression?

- a. Classification can have a numeric or categorical target, while regression has always a categorical target
- b. Classification is a supervised activity, while regression in unsupervised
- c. Classification can make errors, while regression is always exact
- d. Classification has a categorical target, while regression has a numeric target

Annulla la scelta

Domanda **15**

Risposta non ancora data

Punteggio max.: 1,00

Which of the following *is not* a strength point of *Dbscan* with respect to *K-means*

Scegli un'alternativa:

- a. The *effectiveness* even if there are clusters with non-convex shape
- b. The efficiency even in large datasets
- c. The *robustness* with respect to the number of attributes
- d. The *effectiveness*, even in presence of *noise*

Annulla la scelta